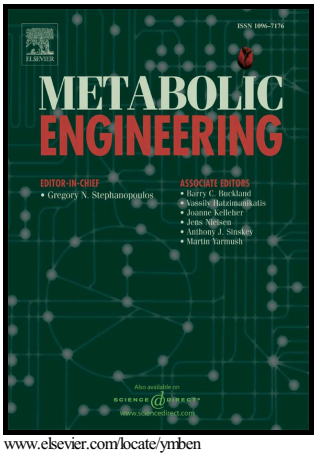


www.elsevier.com/locate/ymben

Production of a bioactive unnatural ginsenoside by metabolically engineered yeasts based on a new UDP-glycosyltransferase from *Bacillus subtilis*

Huichao Liang, Zongfeng Hu, Tingting Zhang, Ting Gong, Jingjing Chen, Ping Zhu, Yan Li, Jinling Yang

PII: S1096-7176(17)30113-1
DOI: http://dx.doi.org/10.1016/j.ymben.2017.07.008
Reference: YMBEN1273

To appear in: *Metabolic Engineering*

Received date: 11 April 2017
Revised date: 5 July 2017
Accepted date: 28 July 2017

Cite this article as: Huichao Liang, Zongfeng Hu, Tingting Zhang, Ting Gong Jingjing Chen, Ping Zhu, Yan Li and Jinling Yang, Production of a bioactive unnatural ginsenoside by metabolically engineered yeasts based on a new UDP glycosyltransferase from *Bacillus subtilis*, *Metabolic Engineering* http://dx.doi.org/10.1016/j.ymben.2017.07.008

Production of a bioactive unnatural ginsenoside by metabolically engineered yeasts based on a new UDP-glycosyltransferase from Bacillus subtilis

Huichao Liang[1], Zongfeng Hu[1], Tingting Zhang, Ting Gong, Jingjing Chen, Ping Zhu, Yan Li, Jinling Yang[*]

State Key Laboratory of Bioactive Substance and Function of Natural Medicines & Key Laboratory of Biosynthesis of Natural Products of National Health and Family Planning Commission, Institute of Materia Medica, Peking Union Medical College & Chinese Academy of Medical Sciences, Beijing 100050, China

[*]Corresponding author. Fax: +86 10 63165199. yangjl@imm.ac.cn

Abstract

Ginsenosides are the main bioactive constituents of Panax species, which are biosynthesized by glycosylation at C3-OH and/or C20-OH of protopanaxadiol (PPD), C6-OH and/or C20-OH of protopanaxatriol (PPT). The C12-glycosylated ginsenosides have scarcely been identified from Panax species. The C12-glycosylated ginsenosides produced from PPD by chemical semi-synthesis have been reported to exhibit higher cytotoxicity than the natural ginsenosides. However, the chemical semi-synthesis approach is not practical due to its complexity and high cost. In our study, a new UDP-glycosyltransferase UGT109A1 was identified from Bacillus subtilis. This enzyme transferred a glucose moiety to C3-OH and C20-OH of dammarenediol-II (DM), C3-OH and C12-OH of PPD and PPT respectively to produce the unnatural ginsenosides 3β-O-Glc-DM, 3β,20S-Di-O-Glc-DM, 3β,12β-Di-O-Glc-PPD and 3β,12β-Di-O-Glc-PPT. Among these unnatural ginsenosides, 3β,12β-Di-O-Glc-PPT is a new compound which has never been reported before. The anti-cancer

[1] These authors contributed equally to this work.

activities of these unnatural ginsenosides were evaluated in vitro and in vivo. 3β,12β-Di-O-Glc-PPD exhibited higher anti-lung cancer activity than Rg3, which is the most active natural ginsenoside against lung cancer. Finally, we constructed metabolically engineered yeasts to produce 3β,12β-Di-O-Glc-PPD by introducing the genes encoding B. subtilis UGT109A1, Panax ginseng dammarenediol-II synthase (DS), P. ginseng cytochrome P450-type protopanaxadiol synthase (PPDS) together with Arabidopsis thaliana NADPH-cytochrome P450 reductase (ATR1) into Saccharomyces cerevisiae INVSc1. The yield of 3β,12β-Di-O-Glc-PPD was increased from 6.17 mg/L to 9.05 mg/L by overexpressing tHMG1. Thus, this study has established an alternative route to produce the unnatural ginsenoside 3β,12β-Di-O-Glc-PPD by synthetic biology strategies, which provides a promising candidate for anti-cancer drug discovery.

Keywords: UDP-glycosyltransferase; Bacillus subtilis; ginsenoside biosynthesis; metabolic engineering; Saccharomyces cerevisiae

1. Introduction

Ginsenosides, a group of triterpenoids saponins, are the main bioactive constituents of Panax species, among which P. ginseng has been used as a traditional herbal medicine in Asia for thousands of years (Shibata, 2001; Yun, 2001). The chemical structures and bioactivities of ginsenosides have been widely studied. Dammarene-type tetracyclic triterpenoids are the major ginsenoside constituents, while Ro is the only oleanane-type pentacyclic ginsenoside, which is found minimally in P. ginseng (Han et al., 2011). Until now, more than 110 naturally occurring dammarane-type ginsenosides have been

identified from Panax species (Christensen, 2009; Yang et al., 2014). Based on the aglycon structures, dammarane-type ginsenosides are mainly divided into two groups. One group are protopanaxadiol (PPD)-type saponins, which are biosynthesized by glycosylation of PPD at its C3-OH and/or C20-OH; the other group are protopanaxatriol (PPT)-type saponins, which are biosynthesized by glycosylation of PPT at its C6-OH and/or C20-OH (Fig. S1). These ginsenosides exhibit a range of bioactivities including anti-cancer, anti-inflammatory, anti-oxidation and neuroprotective effects (Shin et al., 2015). The differences in the position, number and type of sugar moieties lead to the changes in bioactivities and medicinal properties of ginsenosides. For example, Rg3 exhibits anti-cancer activity while Rb1 possesses the influence on the central nervous system (Chen et al., 2008; Xue et al., 2006).

The ginsenoside aglycons also contain a hydroxyl at the C12 position, while the C12-glycosylated ginsenosides have scarcely been identified from Panax species. The C12-glycosylated ginsenosides produced from PPD by chemical semi-synthesis have been reported to exhibit higher cytotoxicity than the natural ginsenosides (Atopkina et al., 1999). However, chemical semi-synthesis faces several challenges including regioselectivity, stereoselectivity, and the protection and de-protection of functional groups. Therefore, large-scale production of C12-glycosylated ginsenosides by chemical semi-synthesis is not practical due to its complexity and high cost (Atopkina et al., 1988; Yu et al., 2013), which limits further exploration of these unnatural ginsenosides in drug discovery.

Currently, metabolic engineering of Saccharomyces cerevisiae has been recognized

as an alternative and promising approach for high-efficient production of rare ginsenosides, such as PPD (Dai et al., 2013), PPT (Dai et al., 2014), Compound K (Yan et al., 2014), Rg3 (Jung et al., 2014; Wang et al., 2015), Rh2 (Wang et al., 2015; Zhuang et al., 2017), F1 and Rh1 (Wei et al., 2015). The biosynthetic pathways of the ginsenoside aglycons PPD and PPT have been elucidated. Triterpenoids are derived from two common building blocks, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) synthesized through mevalonic acid (MVA) pathway in S. cerevisiae. The precursor 2,3-oxidosqualene is primarily cyclized by dammarenediol-II synthase (DS) to form dammarenediol-II (DM) (Tansakul et al., 2006). Subsequently, DM is hydroxylated by PPD synthase (PPDS) at its C12 position to form PPD, and PPD is further hydroxylated by PPT synthase (PPTS) at its C6 position to form PPT (Han et al., 2011; Han et al., 2012). UDP-glycosyltransferases (UGTs), which transfer a sugar from UDP-sugar to aglycons, play an important role in producing different ginsenosides. In the past few years, several UGTs have been identified from P. ginseng. UGTPg1 was reported to regiospecifically glycosylate C20-OH of DM and PPD (Yan et al., 2014). PgUGT74AE2 and PgUGT94Q2 were reported to catalyze the glycosylation of C3-OH of PPD to generate Rh2 and to elongate a glucose moiety of Rh2 to generate Rg3, respectively (Wang et al., 2015). UGTPg100 was reported to catalyze the glycosylation of C6-OH of PPT to produce Rh1 (Wei et al., 2015). However, no UGT has ever been reported to glycosylate C12-OH of PPD or PPT to produce C12-glycosylated ginsenosides in Panax species.

Microorganisms contain rich UGTs capable of using UDP-sugar as the sugar donor

and small molecules as the sugar acceptors (Pandey et al., 2014; Parajuli et al., 2014). Several UGTs have been identified from Bacillus species. Bacillus subtilis CGMCC No.1318 was reported to transform ginsenoside Rh1 into an unnatural ginsenoside 3-O-β-D-glucopyranosyl-6-O-β-D-glucopyranosyl-20(S)-protopanaxatriol (Luo et al., 2013) and the UGT YjiC1 was found responsible for this reaction. This enzyme was cloned and expressed in Escherichia coli and characterized (Luo at al., 2015). BcGT-1 from B. cereus 10987 was reported to glycosylate flavonoids to produce flavonoid glucosides (Jae et al., 2006). The UGT YjiC from B. licheniformis DSM-13 was used to glycosylate phloretin, apigenin, resveratrol and emodin to improve their solubilities and bioactivities (Gurung et al., 2013; Pandey et al., 2014).

In this study, we report a new UDP-glycosyltransferase UGT109A1 (named by the UGT Nomenclature Committee, http://prime.vetmed.wsu.edu/resources/udp-glucuronsyltransferase-homepage) from B. subtilis CTCC 63501, which catalyzes the glycosylation of ginsenoside aglycons at C3-OH, C12-OH or C20-OH to produce unnatural ginsenosides. The recombinant UGT109A1 was expressed in E. coli and characterized. The catalytic properties and enzyme kinetics of UGT109A1 were investigated. The anti-cancer activities of the unnatural ginsenosides produced by UGT109A1 catalysis were evaluated in vitro and in vivo. Finally, we engineered the biosynthetic pathway of the most active unnatural ginsenoside 3,12-Di-O-β-D-glucopyranosyl-dammar-24-ene-3β,12β,20S-triol (3β,12β-Di-O-Glc-PPD) in S. cerevisiae INVSc1 by introducing UGT109A1 and the genes involved in ginsenoside biosynthesis. This study established an alternative route to

produce 3β,12β-Di-O-Glc-PPD by synthetic biology approach.

2. Materials and methods

2.1. Cloning of UDP-glycosyltransferase gene UGT109A1

The genomic DNA was isolated from B. subtilis CTCC 63501 using TIANamp Bacteria DNA Kit (TianGen, Beijing, China). The gene encoding UDP-glycosyltransferase was amplified by PCR using the genomic DNA as the template and the following primers which were designed according to the genomic sequences of B. subtilis 168 (Barbe et al.， 2009): Up-F (5´-ATAAGGAGACTGGAGATTC-3´) and Down-R (5´-ATTGATTTCGGTTTTTATG-3´). The amplified fragment was cloned into pEASY-Blunt-simple vector (Transgen, Beijing, China) to obtain plasmid pEASY-Blunt-UGT109A1. It was then sequenced and the corresponding GenBank accession number was KY952161.

2.2. Heterologous expression of UGT109A1

The gene UGT109A1 was amplified using UGT109A1-F1 and UGT109A1-R1 as primers and pEASY-Blunt-UGT109A as template (Table S1). The amplified fragment was digested with BamH I and Sal I and then cloned into pET32a (Novagen) to obtain the recombinant plasmid pET32a-UGT109A1. The plasmid was then transformed into E. coli BL21 (DE3). The recombinant strains were cultivated in LB medium with 100 μg/mL ampicillin at 37 ˚C, 200 rpm until OD_{600} reached 0.6-0.8. Then UGT109A1 was expressed under 1 mM IPTG induction at 16 ˚C and 180 rpm for 18 h. The cells were collected by centrifugation (8000 g, 5 min), suspended in 100 mM Tris-HCl buffer (pH

8.0) with 1 mM PMSF and disrupted by ultrasonic fragmentation. The cell debris was removed by centrifugation (10000 g, 10 min) and the supernatant was used as crude enzyme for enzymatic assays. The E. coli BL21 (DE3) strains transformed with pET32a were treated in parallel as a control. The recombinant UGT109A1 in the supernatant was purified by Ni-NTA affinity chromatography and then confirmed by SDS-PAGE.

2.3. Activity analysis of recombinant UGT109A1

The reactions were carried out in a total volume of 100 μL containing 20 mM Tris-HCl (pH 8.0), 0.5 mM substrate 20(S)-DM, 20(S)-PPD or 20(S)-PPT, 5 mM UDP-glucose and the crude recombinant UGT109A1. At the same time, a control was performed with the reaction mixture added by the lysate of induced E. coli harboring pET32a. After an incubation period of 12 h or 24 h at 37 ˚C, the reactions were terminated by adding 100 μL methanol. The products were extracted and evaporated. The residue was dissolved in methanol for HPLC and HPLC-ESI-MS analysis.

2.4. Enzymatic assays for UGT109A1

For the kinetic study of UGT109A1, six ginsenosides with different positions of free hydroxyls were selected as the substrates of UGT109A1. Among these ginsenosides, except that Rh2 was commercially available from Nanjing Spring & Autumn Biological Engineering Co. Ltd., the others were all prepared by in vitro enzymatic reactions. The recombinant UGT109A1 expressed in E. coli converted DM, PPD and PPT into 3-O-β-D-glucopyranosyl-dammar-24-ene-3β,20S-diol (3β-O-Glc-DM), 3β,12β-Di-O-Glc-PPD and

3,12-Di-O-β-D-glucopyranosyl-dammar-24-ene-3β,6α,12β,20S-tetraol (3β,12β-Di-O-Glc-PPT), respectively. Then 12-O-β-D-glucopyranosyl-dammar-24-ene-3β,12β,20S-triol (12β-O-Glc-PPD) and 12-O-β-D-glucopyranosyl-dammar-24-ene-3β,6α,12β,20S-tetraol (12β-O-Glc-PPT) were produced by the catalysis of recombinant glycoside hydrolase LXYL-P1-2 from the fungus Lentinula edodes expressed in E. coli (Cheng et al., 2013), which selectively hydrolyzed the C3 glycosyl of both 3β,12β-Di-O-Glc-PPD and 3β,12β-Di-O-Glc-PPT. In addition, 20-O-β-D-glucopyranosyl-dammar-24-ene-3β,20S-diol (20S-O-Glc-DM) was prepared by in vitro enzymatic reaction catalyzed by the recombinant UGTPg1 expressed in E. coli (Yan et al., 2014). The recombinant PgUGT74AE2 (Wang et al., 2015) catalyzed the C3-OH of PPT to produce 3-O-β-D-glucopyranosyl-dammar-24-ene-3β,6α,12β,20S-tetraol (3β-O-Glc-PPT). All the ginsenosides prepared by enzymatic reactions were confirmed by HPLC-ESI-MS, ^{1}H-NMR, ^{13}C-NMR and HMBC analysis.

The catalytic activity of UGT109A1 was investigated in the reaction mixture of 100 μL containing 20 mM Tris-HCl buffer, 0.5 mM substrate (20S-O-Glc-DM, 12β-O-Glc-PPD, 12β-O-Glc-PPT, Rh2, 3β-O-Glc-PPT or 3β-O-Glc-DM), 5 mM UDP-glucose and 500 ng purified UGT109A1. The reactions were terminated by adding 100 μL methanol. The products were extracted and evaporated. HPLC analysis was used to quantify the target products in each reaction. The effect of temperature on enzymatic activity was tested at varied temperatures ranging from 10 ˚C to 60 ˚C in 20 mM Tris-HCl (pH 8.0) for 2 h. The optimal pH was tested at 37 ˚C in each buffer with varied

pH ranging from 5.0 to 12.0 for 2 h. All data are presented as means±SD of three independent repeats.

The kinetic parameters of UGT109A1 were determined in the reaction mixture contained 20 mM Tris-HCl buffer, substrate (0.025, 0.05, 0.1, 0.2, 0.4, 0.8, 1.6 mM) , 5 mM UDP-glucose and 500 ng purified UGT109A1 in a final volume of 100 μL. The reactions were incubated at optimal temperature and pH with each substrate for 2 h. Finally, the reactions were terminated by adding 100 μL methanol. The products were extracted and evaporated. HPLC analysis was used to quantify the target product in each reaction. The Michaelis-Menten parameters were calculated by least-squares fitting of the kinetic model using GraphPad Prism 5. All data are presented as means±SD of three independent repeats.

2.5. In vitro cytotoxicity test of unnatural ginsenosides

The anti-cancer activities of unnatural ginsenosides (purity >98%) produced by UGT109A1 catalysis were determined by 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) method with 20(S)-Rg3 (purity >98%, purchased from Nanjing Spring & Autumn Biological Engineering Company) as the positive control. The cancer cell lines included human breast cancer cells (MCF-7), human hepatoma carcinoma cells (HepG2), human lung cancer cells (NCI-H460) and human pancreatic cancer cells (Capan2). Cell lines were seeded in a 96-well plate. After incubation for 24 h, cells were treated with varied concentrations of unnatural ginsenosides or dimethyl sulfoxide as vehicle. After incubation for 120 h, 50 μL MTT stock solution (2 mg/mL, Sigma Chemical) was added to each well, and the plates were incubated for an additional 4 h at 37 ˚C. The solution

was then removed from each well, and 150 μL dimethyl sulfoxide was added. Following gentle agitation, the absorbance was measured using an ELISA reader (Bio-Rad, USA) at 570 nm. Three parallel samples were measured each time. Absorbance values were normalized to the values obtained for vehicle-treated cells to determine the percentage of surviving cells. The median inhibitory concentration (IC_{50}) was assessed from the dose response curve.

2.6. In vivo evaluation of anti-lung cancer activity of 3β,12β-Di-O-Glc-PPD

C57BL/c mice (males, 6-8 weeks old) were purchased from Vital River Laboratory Animal Technology Co., Ltd. (Beijing, China) and housed in controlled environment at 25 ˚C on a 12 h light/dark cycle. All animal protocols were conformed to the Guidelines for the Care and Use of Laboratory Animals approved by the Animal Care and Use Committee of Chinese Academy of Medical Sciences and Peking Union Medical College. Lewis (Murine lung cancer) cells were harvested and washed three times with normal saline. The cells were counted and diluted to 5×10^7 cells/mL. C57BL/c mice were subcutaneously implanted with 0.2 mL cell dilution on the right flank. 24 h after inoculation, C57BL/c mice were randomly divided into 12 groups. One group received p.o of 25% PEG400 as a model control. One group received injection of 15.0 mg/kg of Taxol (twice every week). 5 groups involved 10.0 mg/kg 20(S)-Rg3 used as the positive control and 2.5 mg/kg, 5.0 mg/kg, 10.0 mg/kg, 20.0 mg/kg 3β,12β-Di-O-Glc-PPD (dissolved in 25% PEG400) were continuously administrated p.o. for 11 days, once a day. 5 groups were administrated 10.0 mg/kg 20(S)-Rg3 or 2.5 mg/kg, 5.0 mg/kg, 10.0 mg/kg, 20.0 mg/kg 3β,12β-Di-O-Glc-PPD with 15.0 mg/kg Taxol (twice every week) simultaneously. At the end of the treatment period, the mice were euthanized and the

tumors were excised and weighed. The inhibition rate (IR) of tumor growth was calculated by the following formula: IR (%) = [(A-B)/A]/100, where A is the average tumor weight of model control, B is the average tumor weight of compound groups.

2.7. Construction of recombinant expression vectors

P. ginseng callus and Arabidopsis thaliana leaves were induced by methyl jasmonate (MeJA). Total RNA was isolated from the induced materials by using RNApure Plant Kit (ComWin Biotech, Beijing, China) and the cDNA was prepared using HiFiScript cDNA Synthesis Kit (ComWin Biotech, Beijing, China). The genes of DS and PPDS were amplified from P. ginseng cDNA by PCR, while NADPH-cytochrome P450 reductase gene ATR1 was amplified from A. thaliana cDNA. The genomic DNA of S. cerevisiae INVSc1 was isolated using TIANamp Yeast DNA Kit (TianGen, Beijing, China). The N-terminally truncated 3-hydroxy-3-methylglutaryl-CoA (HMG-CoA) reductase gene tHMG1 and transcription factor gene UPC2.1 were amplified from yeast genomic DNA. All the primers used for gene cloning in this study are listed in Table S1. All the amplified fragments were cloned into pEASY-Blunt-simple vector and then sequenced.

DS and green fluorescent protein gene GFP were used as the templates and amplified to produce DS-GFP fragment by PCR (Liang et al., 2016). The fragment DS-GFP was designed with EcoR I and Sac I and then inserted into pESC-HIS to construct pESC-HIS-DS-GFP. The fragment ATR1 was designed with BamH I and Sal I and then inserted into pESC-TRP to construct pESC-TRP-ATR1. The fragment UPC2.1 was designed with Not I and Sac I and then inserted into pESC-TRP-ATR1 to construct pESC-TRP-ATR1-UPC2.1. The fragment PPDS was designed with BamH I and Sal I

and then inserted into pESC-URA-ERG7^{-} (a recombinant expression vector with antisense lanosterol synthase gene constructed in our previous study) (Wang et al., 2015) to construct pESC-URA-ERG7^{-}-PPDS. The fragment UGT109A1 was designed with Not I and Spe I and then inserted into pESC-LEU to construct pESC-LEU-UGT109A1. The fragment tHMG1 was designed with BamH I and Sal I and then inserted into pESC-LEU-UGT109A1 to construct pESC-LEU-tHMG1-UGT109A1. All the recombinant plasmids constructed in this study are listed in Table 1.

2.8. Construction of engineered yeasts

Transformation of S. cerevisiae strains was performed by standard lithium acetate method (Gietz et al., 1995). The four plasmids pESC-HIS-DS-GFP, pESC-TRP-ATR1, pESC-URA-ERG7^{-}-PPDS and pESC-LEU-UGT109A1 were co-transformed into INVSc1 and the resulted recombinant was named as IN-A. The four plasmids pESC-HIS-DS-GFP, pESC-TRP-ATR1, pESC-URA-ERG7^{-}-PPDS and pESC-LEU-tHMG1-UGT109A1 were co-transformed into INVSc1 and the resulted recombinant was named as IN-B. The four plasmids pESC-HIS-DS-GFP, pESC-TRP-ATR1-UPC2.1, pESC-URA-ERG7^{-}-PPDS and pESC-LEU-tHMG1-UGT109A1 were co-transformed into INVSc1 and the resulted recombinant was named as IN-C. All the recombinant strains were verified by PCR and listed in Table 1.

2.9. Yeast cultivation and metabolite extraction

To determine the products of the engineered yeasts, individual clones were

inoculated into the SD medium lacking the corresponding nutrition (2% glucose as the sole carbon source) and cultivated at 30 ˚C, 200 rpm for 24 h. Aliquots were diluted to an initial OD_{600} of 0.8 in 100 mL of the SG medium lacking the corresponding nutrition (2% galactose as sole carbon source) in 500 mL flasks and grown at 30 ˚C, 200 rpm for 10 days. The cells were harvested by centrifugation (5000 g, 10 min), freeze-dried by vacuum freeze-dryer and weighed. Then cell pellets were disrupted and extracted with n-butanol and used for the analysis of intracellular 3β,12β-Di-O-Glc-PPD. The supernatant was extracted directly with n-butanol and used for the analysis of extracellular 3β,12β-Di-O-Glc-PPD. All data are presented as means ±SD of three independent repeats.

Table 1 Plasmids and strains used in this study.

	Genotype or characteristic	Source
Plasmids		
pESC-HIS-DS-GFP	P_{GAL10}-DS-GFP-T_{ADH1}	This work
pESC-URA-ERG7^{-}	P_{GAL10}-ERG7^{-}-T_{ADH1}	Wang et al., 2015
pESC-URA-ERG7^{-}-PPDS	P_{GAL1}-PPDS-T_{CYC1}; P_{GAL10}-ERG7^{-}-T_{ADH1}	This work
pESC-TRP-ATR1	P_{GAL1}-ATR1-T_{CYC1}	This work
pESC-TRP-ATR1-UPC2.1	P_{GAL1}-ATR1-T_{CYC1}; P_{GAL10}-UPC2.1-T_{ADH1}	This work
pESC-LEU-UGT109A1	P_{GAL10}-UGT109A1-T_{ADH1}	This work
pESC-LEU-tHMG1-UGT109A1	P_{GAL1}-tHMG1-T_{CYC1}; P_{GAL10}-UGT109A1-T_{ADH1}	This work
Strains		
INVSc1	MATa, his3Δ1, leu2, trp1-289, ura3-52/MATα, his3Δ1, leu2, trp1-289, ura3-52	Invitrogen
IN-A	pESC-HIS-DS-GFP; pESC-TRP-ATR1; pESC-URA-ERG7^{-}-PPDS; pESC-LEU-UGT109A1	This work
IN-B	pESC-HIS-DS-GFP; pESC-TRP-ATR1; pESC-URA-ERG7^{-}-PPDS; pESC-LEU-tHMG1-UGT109A1	This work
IN-C	pESC-HIS-DS-GFP; pESC-TRP-ATR1-UPC2.1; pESC-URA-ERG7^{-}-PPDS; pESC-LEU-tHMG1-UGT109A1	This work

2.10. HPLC, HPLC-ESI-MS and NMR analysis

HPLC analysis was performed on an Agilent 1200 system equipped with a binary

pump, an online degasser, an autoplate-sampler and a thermostatically controlled column compartment. Chromatographic separation was carried out at 28 ˚C on a Cosmosil 5C_{18}-MS-II column (5 μm, 4.6 mm×150 mm). Semi-preparative HPLC was performed on a HPLC system equipped with a Shimadzu LC-6AD pump and a Shimadzu SPD-20A prominence UV-VIS detector (Shimadzu Corporation, Kyoto, Japan) using an Agilent C_{18} column (5 μm, 9.6 mm×250 mm). The gradient elution system consisted of water and acetonitrile. The detailed conditions of analytic and preparative HPLC for extracts of enzyme reaction and yeast fermentation are described in Table S2.

Chromatographic analysis and ESI-MS data were acquired using HPLC-LTQ/FTICR-MS system (Thermo Fisher Scientific, Massachusetts, United States) equipped with a Cosmosil 5C_{18}-MS-II column (5 μm, 4.6 mm×150 mm). NMR experiments were performed in CD_3OD for all ginsenosides on INOVA 500 and 600 NMR spectrometer (Varian, California, United States) referenced to the solvent peaks.

3. Results and discussion

3.1. Cloning and heterologous expression of UGT109A1

The UDP-glycosyltransferase gene was amplified from the total DNA of B. subtilis CTCC 63501 by PCR using primers based on the published sequences and designated as UGT109A1. The sequence of UGT109A1 was 1179 bp in length (Fig. S2), encoding a deduced amino acid sequence with 94.9% identity to that of YjiC1, the first characterized glycosyltransferase from B. subtilis which could transfer a glucose moiety to the free C3-OH of ginsenoside Rh1 (Luo at al., 2015). The gene was cloned into pET32a and

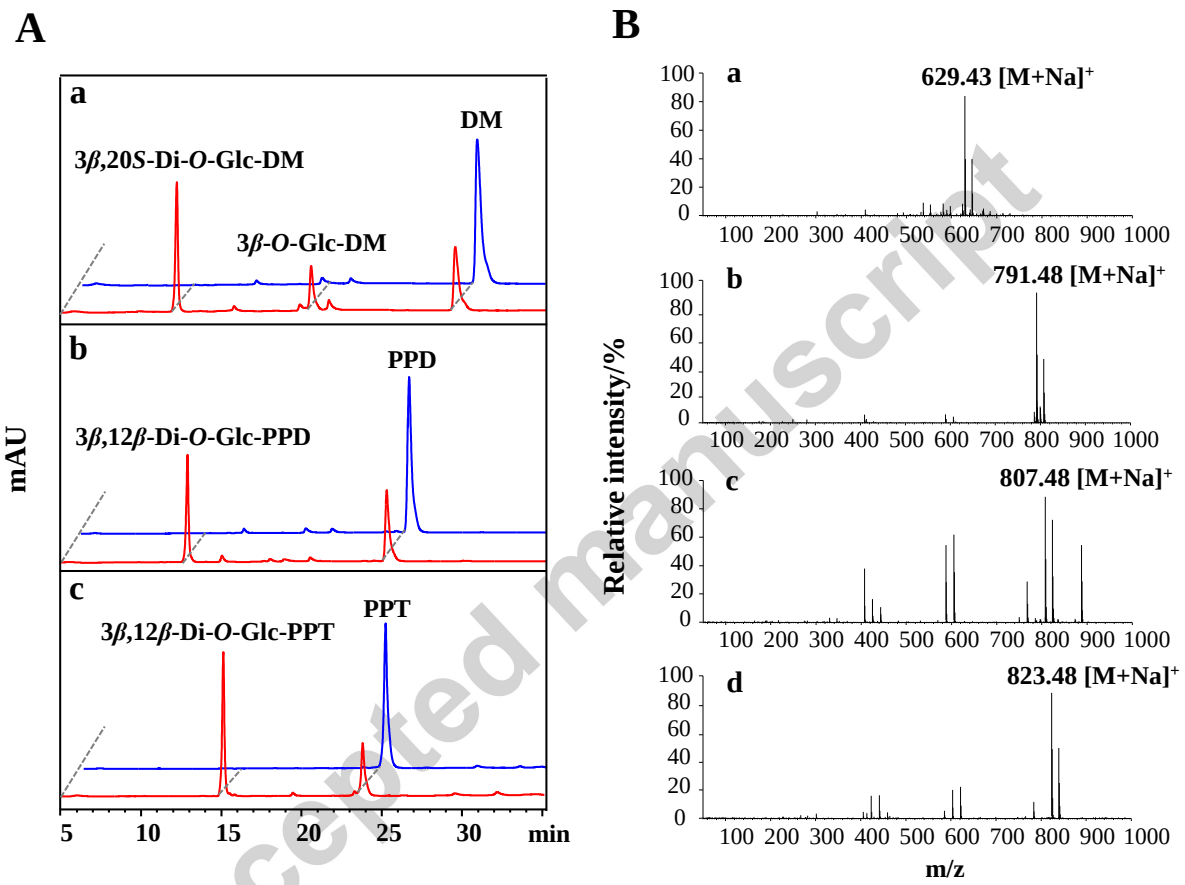
heterologously expressed in E. coli BL21 (DE3). SDS-PAGE analysis showed that the recombinant UGT109A1 was expressed in soluble form (Fig. S3). The recombinant UGT109A1 was purified by Ni-NTA affinity chromatography and then confirmed by SDS-PAGE (Fig. S4).

3.2. Enzymatic activities of UGT109A1 against DM, PPD and PPT

In order to identify the enzymatic activity of UGT109A1, the crude cell extract containing the recombinant UGT109A1 was incubated with UDP-glucose as sugar donor and DM, PPD and PPT as sugar acceptors. After an incubation period of 24 h at 37 ˚C, the recombinant UGT109A1 transferred a glucose moiety to both the free C3-OH and C20-OH of DM to yield 3β-O-Glc-DM and 3,20-Di-O-β-D-glucopyranosyl-dammar-24-ene-3β,20S-diol (3β,20S-Di-O-Glc-DM), which were confirmed by HPLC-ESI-MS, ^{1}H-NMR, ^{13}C-NMR and HMBC (Fig. 1A-C, Figs. S5-6). Similarly, UGT109A1 transferred a glucose moiety to both the free C3-OH and C12-OH of PPD or PPT to yield 3β,12β-Di-O-Glc-PPD or 3β,12β-Di-O-Glc-PPT (Fig. 1A-C, Figs. S7-8). When the incubation period was shortened to 12 h, Rh2 or 3β-O-Glc-PPT was also detected respectively as an intermediate in the reaction systems (Fig. S9). However, neither C20-glycosylated PPD-type nor C20-glycosylated PPT-type ginsenoside was detected even if the reaction conditions were adjusted. To our knowledge, this is the first report of UGT to glycosylate C12-OH of PPD and PPT so far. Moreover, all the above mentioned products except Rh2 generated by UGT109A1 catalysis were unnatural ginsenosides which have never been identified from Panax species. In particular, it is worth mentioning that 3β,12β-Di-O-Glc-PPT was a new

compound which has never been reported before (Fig. 1C).

Other UDP-activated sugar donors such as UDP-galactose, UDP-glucuronic acid and UDP-N-acetylglucosamine were also used to incubate with the recombinant UGT109A1 and DM, PPD or PPT, but no products were detected. The results indicated that UGT109A1 had a relatively strict specificity for sugar donors.



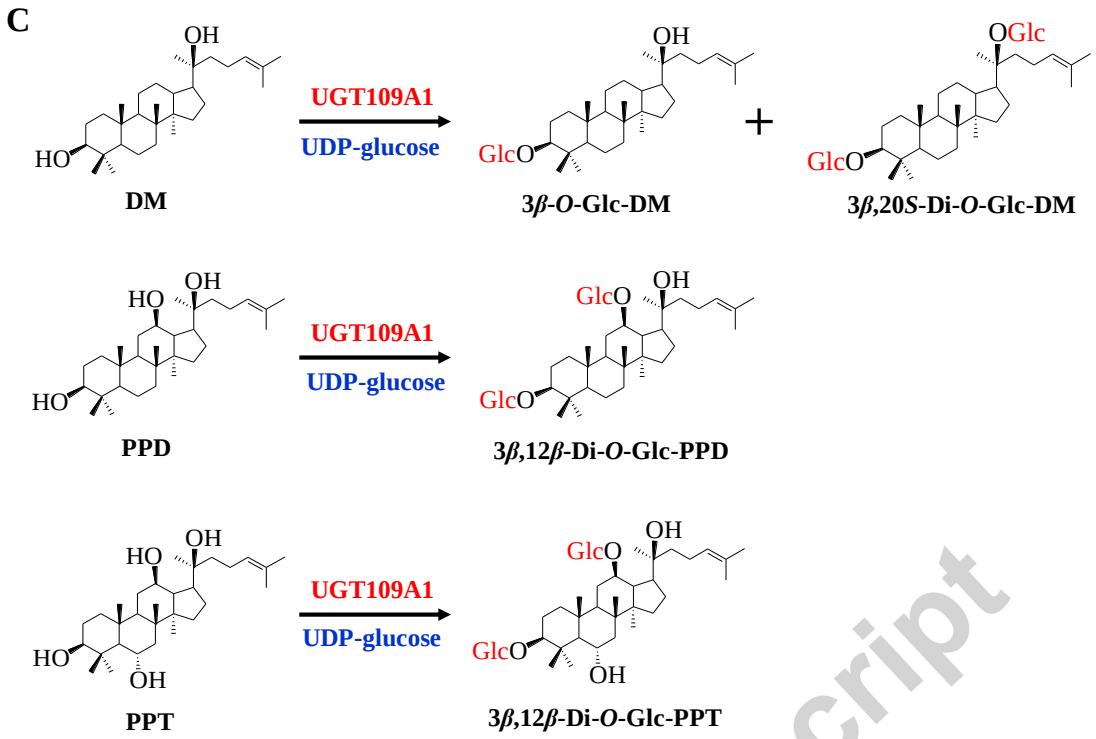


Figure 1. Functional characterization of the UDP-glycosyltransferase UGT109A1 from B. subtilis. (A) HPLC analysis of the reaction products from the incubation of UGT109A1 with DM, PPD or PPT as substrates for 24 h. Red, the cell extract of recombinant E. coli harboring pET32a-UGT109A1; Blue, negative control, the cell extract of recombinant E. coli harboring pET32a; (a) The reaction of UGT109A1 catalyzing DM; (b) The reaction of UGT109A1 catalyzing PPD; (c) The reaction of UGT109A1 catalyzing PPT. (B) The MS spectrum of the products from the reactions of UGT109A1 catalyzing DM, PPD or PPT. (a) The MS spectrum of 3β-O-Glc-DM; (b) The MS spectrum of 3β,20S-Di-O-Glc-DM; (c) The MS spectrum of 3β,12β-Di-O-Glc-PPD; (d) The MS spectrum of 3β,12β-Di-O-Glc-PPT. (C) The schematic presentation of unnatural ginsenosides production based on the function of UGT109A1.

3.3. Catalytic property and kinetic study of UGT109A1

To investigate the enzymatic property of UGT109A1, six ginsenosides

(20S-O-Glc-DM, 12β-O-Glc-PPD, 12β-O-Glc-PPT, Rh2, 3β-O-Glc-PPT and 3β-O-Glc-DM) with free hydroxyls at different positions were selected as the substrates of UGT109A1 (Fig. 2). Among these substrates, except that Rh2 was commercially available, the others were all prepared by in vitro enzymatic reactions. All the substrates prepared by enzymatic reactions were confirmed by HPLC-ESI-MS, ^{1}H-NMR, ^{13}C-NMR and HMBC analysis, which were described in detail in Supplementary Data (Figs. S5, S10-13).

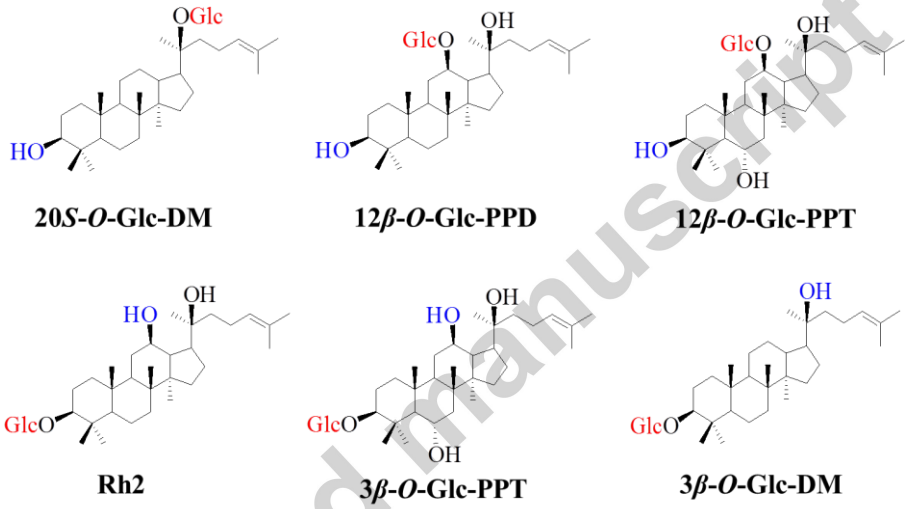


Figure 2. The structures of the substrates used for enzymatic assays of UGT109A1

The reactions were conducted at different temperatures (Fig. S14) and pH (Fig. S15), and the optimal temperature and pH of the reactions for each substrate catalyzed by UGT109A1 were listed in Table 2. As shown in Table 2, UGT109A1 possessed the maximal activity in the pH range of 9.0-10.0 and in the temperature range of 35 ˚C-40 ˚C.

The kinetic parameters for UGT109A1 to glycosylate the free hydroxyls at different positions of substrates were determined at the optimal temperature and pH. The results

were listed in Table 3 and Fig. S16. The glycosylated positions of 12β-O-Glc-PPD and 12β-O-Glc-PPT catalyzed by UGT109A1 were both C3-OH, but the K_m of UGT109A1 against 12β-O-Glc-PPT was approximately 5-fold higher than that of UGT109A1 against 12β-O-Glc-PPD, indicating that the affinity of UGT109A1 against 12β-O-Glc-PPD is much higher than that of UGT109A1 against 12β-O-Glc-PPT. Meanwhile, the turnover number (k_{cat}) for 12β-O-Glc-PPD was about 10-fold higher than that of 12β-O-Glc-PPT. Consequently, the catalytic efficiency (k_{cat}/K_m) of UGT109A1 against 12β-O-Glc-PPD was almost 50-fold higher than that of UGT109A1 against 12β-O-Glc-PPT. Moreover, the turnover number (k_{cat}) of UGT109A1 against 3β-O-Glc-PPT was one half of that of UGT109A1 against Rh2, but the affinity (K_m) of the former was nearly equal to that of the latter, suggesting that UGT109A1 catalyzed the glycosylation of C12-OH of PPD-type ginsenosides with higher efficiency than that of PPT-type ginsenosides. These results indicated that the existence of C6-OH reduced the glycosylation efficiency of C3-OH and C12-OH by UGT109A1. In particular, the glycosylation of C3-OH was more repressed by C6-OH than that of C12-OH, which might be due to the fact that C3-OH is nearer to C6-OH in distance at the spatial level. In addition, the affinity of UGT109A1 against 3β-O-Glc-DM was 2-fold higher than that of UGT109A1 against Rh2 and 3β-O-Glc-PPT, while the k_{cat} of UGT109A1 against Rh2 and 3β-O-Glc-PPT was 65 and 32-fold higher than that of UGT109A1 against 3β-O-Glc-DM respectively, indicating that UGT109A1 was more efficient in catalyzing the C12-OH glycosylation of Rh2 and 3β-O-Glc-PPT than the C20-OH glycosylation of 3β-O-Glc-DM. This finding may account for the results of the reactions, in which no Rh2 or 3β-O-Glc-PPT or

C20-glycosylated products were accumulated when UGT109A1 catalyzed PPD or PPT for 24 h, but 3β-O-Glc-DM was accumulated when UGT109A1 catalyzed DM (Fig. 1; Table 3).

Table 2 The optimal catalysis conditions of UGT109A1 against different substrates.

Substrate	Glycosylation site	Optimal pH	Optimal temperature
20S-O-Glc-DM	C3	10.0	40 °C
12β-O-Glc-PPD	C3	9.0	35 °C
12β-O-Glc-PPT	C3	9.0	40 °C
Rh2	C12	9.0	35 °C
3β-O-Glc-PPT	C12	9.0	35 °C
3β-O-Glc-DM	C20	10.0	40 °C

Table 3 The kinetic parameters of UGT109A1 against different substrates.

Substrate	Glycosylation site	V_{max} (nmol/min/μg)	K_m (mM)	k_{cat} (S^{-1})	k_{cat}/K_m ($mM^{-1}\cdot S^{-1}$)
20S-O-Glc-DM	C3	0.499(±0.021)	0.343(±0.039)	0.512(±0.022)	1.490(±0.030)
12β-O-Glc-PPD	C3	0.434(±0.009)	0.163(±0.011)	0.889(±0.018)	5.461(±0.278)
12β-O-Glc-PPT	C3	0.087(±0.003)	0.767(±0.049)	0.089(±0.003)	0.116(±0.005)
Rh2	C12	1.992(±0.096)	0.402(±0.050)	2.042(±0.098)	5.084(±0.166)
3β-O-Glc-PPT	C12	0.992(±0.056)	0.414(±0.060)	1.017(±0.057)	2.453(±0.028)
3β-O-Glc-DM	C20	0.031(±0.002)	0.203(±0.027)	0.031(±0.002)	0.155(±0.008)

3.4. In vitro cytotoxicity test of unnatural ginsenosides

The unnatural ginsenosides produced by UGT109A1 catalysis were tested for in vitro cytotoxicity against four human cancer cell lines (MCF-7, HepG2, NCI-H460 and Capan2) by MTT assay. Rg3 was used as the positive control because it is the most active natural ginsenoside against lung cancer (Shinkai et al., 1996; Zhang et al., 2006). Among all the tested unnatural ginsenosides, 3β,12β-Di-O-Glc-PPD showed the highest cytotoxicity against lung cancer NCI-H460 cells with the IC_{50} of 84.6 μM, while the IC_{50} of Rg3 was 271.0 μM (Table 4). The result was in agreement with the previous finding

that semi-synthetic unnatural ginsenosides displayed higher cytotoxicity against human lung carcinoma cell lines GLC4 than natural ginsenosides (Atopkina et al., 1999). This result suggested that 3β,12β-Di-O-Glc-PPD may be more effective against lung cancer than Rg3.

Table 4 In vitro cytotoxicity test of the unnatural ginsenosides against cancer cell lines.

Compound	IC_{50} (μM)			
	MCF-7	HepG2	NCI-H460	Capan2
Rg3	356.0	427.0	271.0	>500.0
3β-O-Glc-DM	224.0	239.0	180.0	272.0
3β,20S-Di-O-Glc-DM	157.0	546.0	389.0	>500.0
3β,12β-Di-O-Glc-PPD	387.0	>500.0	84.6	446.0
3β,12β-Di-O-Glc-PPT	>500.0	>500.0	>500.0	>500.0

3.5. In vivo evaluation of anti-lung cancer activity of 3β,12β-Di-O-Glc-PPD

Since 3β,12β-Di-O-Glc-PPD exhibited high cytotoxicity against lung cancer cell lines, we further investigated its anti-lung cancer activity in mice. Lewis lung cancer xenograft mouse model was used to evaluate the in vivo anti-lung cancer activity of 3β,12β-Di-O-Glc-PPD with Rg3 as the positive control (Table 5). The results showed that both Rg3 and 3β,12β-Di-O-Glc-PPD inhibited the tumor growth significantly in Lewis xenografts model. In the group treated with 5 mg/kg 3β,12β-Di-O-Glc-PPD, the inhibition rate of the tumor weight was 49.69%, which was apparently higher than the group treated with 10 mg/kg Rg3 with the inhibition rate of 34.59%. Furthermore, both Rg3 and 3β,12β-Di-O-Glc-PPD markedly increased the inhibition rate of the tumor weight when they were combined with Taxol in treatment. The inhibition rate of 5 mg/kg 3β,12β-Di-O-Glc-PPD plus 15 mg/kg Taxol was 64.64%, which was much higher than that of Taxol alone (15 mg/kg) with the inhibition rate of 44.68%. The synergistic effect was also better than 10 mg/kg Rg3 plus 15 mg/kg Taxol with the inhibition rate of 53.39%. It's encouraging that 3β,12β-Di-O-Glc-PPD displayed higher anti-lung cancer

activity even at a lower concentration than Rg3 whether used alone or combined with Taxol in treatment. These findings suggested that 3β,12β-Di-O-Glc-PPD may be a promising candidate for anti-lung cancer drugs.

Table 5 In vivo evaluation of anti-cancer activity of 3β,12β-Di-O-Glc-PPD in Lewis lung cancer xenografts mouse model.

Group	Dosage (mg/kg)	Schedule of administration	Animal number	Body weight (g)		Tumor weight (g)	Inhibition rate (%)
				Begin	End		
Vehicle control			6/6	17.8±0.8	21.1±1.0	2.16±0.43	
Taxol	15.0	1,4,7	6/6	18.0±0.7	21.1±0.7	1.20±0.22***	44.68
Rg3	10.0	1-11	6/6	18.4±0.8	22.7±1.3	1.42±0.63*	34.59
3β,12β-Di-O-Glc-PPD	2.5	1-11	6/6	17.6±0.8	20.4±1.2	1.91±0.54	11.94
	5.0	1-11	6/6	18.6±0.9	22.8±0.6	1.09±0.60**	49.69
	10.0	1-11	6/6	18.0±0.6	20.6±1.2	1.28±0.32**	40.99
	20.0	1-11	6/6	18.2±0.6	21.4±1.4	1.55±0.41*	28.43
Taxol+Rg3	15.0+10.0	1,4,7+1-11	6/6	18.0±0.6	18.8±1.5	1.01±0.45**	53.39
Taxol+ 3β,12β-Di-O-Glc-PPD	15.0+2.5	1,4,7+1-11	6/6	18.1±0.8	19.3±1.2	1.08±0.36***	50.15
	15.0+5.0	1,4,7+1-11	6/6	18.1±0.6	20.7±0.6	0.77±0.34***Δ	64.64
	15.0+10.0	1,4,7+1-11	6/6	18.0±0.5	18.1±1.6	0.84±0.30***Δ	61.17
	15.0+20.0	1,4,7+1-11	6/6	17.8±0.7	19.3±0.7	1.17±0.17***	46.07

*p<0.05, **p<0.01, *** p<0.001, compared with vehicle control; Δp<0.05, compared with Taxol

3.6. Construction of engineered yeasts for the production of 3β,12β-Di-O-Glc-PPD

PPD is a precursor of dammarane-type ginsenosides and its biosynthetic pathway has been elucidated in the previous studies (Han et al., 2006, 2011; Tansakul et al., 2006). The biosynthetic pathway of PPD has also been constructed successfully in S. cerevisiae BY4742 (Dai et al., 2013). Based on this study, the biosynthetic pathway of 3β,12β-Di-O-Glc-PPD was engineered in the yeast host to produce

3β,12β-Di-O-Glc-PPD effectively. The engineered pathway involved the MVA pathway from glucose to IPP and DMAPP, the pathway from IPP and DMAPP to 2,3-oxidosqualene, and the heterologous pathway from 2,3-oxidosqualene to 3β,12β-Di-O-Glc-PPD (Fig. 3). The pathway from glucose to 2,3-oxidosqualene is intrinsic in the yeast host. In the heterologous pathway, DS and PPDS were cloned from P. ginseng, ATR1 was cloned from A. thaliana, and UGT109A1 was cloned from B. subtilis. DS cyclizes 2,3-oxidosqualene to produce DM (Tansakul et al., 2006), then PPDS together with ATR1 hydroxylates DM to produce PPD (Dai et al., 2013; Wang et al., 2012), and finally UGT109A1 catalyzes PPD to produce 3β,12β-Di-O-Glc-PPD.

In the proposed biosynthetic pathway of 3β,12β-Di-O-Glc-PPD, cyclization of 2,3-oxidosqualene to DM by DS is the first step towards the biosynthesis of 3β,12β-Di-O-Glc-PPD, which is also the key branch point of ergosterol and 3β,12β-Di-O-Glc-PPD biosynthesis in the engineered yeasts. Down-regulation of 2,3-oxidosqualene metabolic flux to ergosterol may redirect more metabolic flux towards the 3β,12β-Di-O-Glc-PPD biosynthetic pathway. In our previous study, the biosynthetic pathway of ergosterol was successfully down-regulated by transforming antisense lanosterol synthase (ERG7) gene in S. cerevisiae INVSc1 (Wang et al., 2015). In another report from our group, DS was fused with the green fluorescent protein (GFP) to study the expression and subcellular localization of recombinant DS in INVSc1. Unexpectedly, the fusion expression of DS with GFP significantly improved the production of dammarenediol-II (Liang et al., 2016). The above findings provided new strategies for the optimization of the biosynthetic pathway of ginsenosides in the engineered yeasts.

Therefore, we constructed the engineered yeasts with fusion expression and antisense technology for the high yield of 3β,12β-Di-O-Glc-PPD.

The recombinant plasmids of pESC-HIS-DS-GFP, pESC-URA-ERG7^{-}-PPDS, pESC-TRP-ATR1 and pESC-LEU-UGT109A1 were introduced into INVSc1 simultaneously and the transformants were selected and verified by PCR. When cultivated in SD medium for 10 days with 2% galactose as both sole carbon source and inducer, HPLC analysis of intracellular and extracellular products indicated that the yields of 3β,12β-Di-O-Glc-PPD varied greatly, suggesting different copies of genes had been introduced into the recombinant strains respectively. The engineered strain with the highest yield of 3β,12β-Di-O-Glc-PPD, a total yield of 6.17 mg/L, was named as IN-A. The structure of 3β,12β-Di-O-Glc-PPD produced by IN-A was confirmed by ^{1}H-NMR, ^{13}C-NMR (Fig. 4A-C, Fig. S17). HMG-CoA reductase is the first rate-limiting enzyme in the MVA pathway, and overexpression of a truncated HMG-CoA reductase gene tHMG1 has been routinely used to increase the carbon flux through the MVA pathway (Kim et al., 2015). This strategy has been successfully adopted to increase the production of taxadiene (Engels et al., 2008), artemisinic acid (Paddon et al., 2013), miltiradiene (Dai et al., 2012) and gisnenosides (Dai et al., 2013; Wang et al., 2015). Thus, in order to enhance the production of 3β,12β-Di-O-Glc-PPD, we also introduced tHMG1 into the yeast host to increase supplies of its precursor 2,3-oxidosqualene, and the transformants were selected subsequently. The engineered strain with the highest yield of 3β,12β-Di-O-Glc-PPD was named as IN-B. The total yield of 3β,12β-Di-O-Glc-PPD in IN-B was 9.05 mg/L, which was 1.47-fold of that produced in IN-A (Fig. 4A-C).

Moreover, UPC2.1, a global transcription factor regulating the biosynthesis of sterols in S. cerevisiae, was proved to increase amorphadiene production when overexpressed in engineered yeast (Ro et al., 2006). In order to further enhance the production of 3β,12β-Di-O-Glc-PPD, UPC2.1 was also introduced into the yeast host, leading to the recombinant stain IN-C. Unexpectedly, no 3β,12β-Di-O-Glc-PPD was detected in IN-C.

The cell biomass and the yields of total dammarane-type triterpenoids (including DM, PPD and 3β,12β-Di-O-Glc-PPD) and the sterols (including lanosterol and ergosterol) in the ergosterol biosynthesis branch were compared among the recombinant stains IN-A, IN-B and IN-C (Table 6). The cell biomass of IN-B after fermentation was slightly lower than that of IN-A, which was almost the same as that of the strain producing PPD (3.71 g/L DCW), suggesting that the production of 3β,12β-Di-O-Glc-PPD did not obviously affect the growth of the engineered strains. In contrast, the cell biomass of IN-C decreased markedly compared with those of IN-A and IN-B due to the interference caused by the overexpression of UPC2.1. The total dammarane-type triterpenoids produced by IN-B was 2-fold higher than that produced by IN-A while IN-B produced only a little more sterols than IN-A. Both squalene and 2,3-oxidosqualene are precursors of both dammarane-type triterpenoids and sterols. The yield of squalene in IN-B was 14-fold higher than that of IN-A, suggesting the enhanced production of the total dammarane-type triterpenoids in IN-B was resulted from the increased supplies of their precursors by overexpression of tHMG1. 2,3-Oxidosqualene was undetected in all engineered strains, which was consistent with the result of previous research (Wang et al., 2015). This was probably because 2,3-oxidosqualene as an intermediate could be rapidly

transformed to PPD and lanosterol. No significant difference in the total yield of dammarane-type triterpenoids between IN-B and IN-C while the yield of sterols in IN-C was almost 2-fold higher than that of IN-B. This observation suggested that the overexpression of UPC2.1 could increase the production of sterols in IN-B while severely inhibited the catalytic activity of UGT109A1. Moreover, although UGT109A1 could catalyze the glycosylation of C3-OH and C20-OH of DM, neither 3β-O-Glc-DM nor 3β,20β-Di-O-Glc-DM was detected in IN-A and IN-B, which was probably because the catalytic efficiency of UGT109A1 against PPD and its derivatives was much higher than against DM and its derivatives (Table 3). Furthermore, mono-glucosylated products including Rh2 and 12β-O-Glc-PPD were undetected in IN-A and IN-B, perhaps due to the fact that UGT109A1 had strong catalytic efficiency of UGT109A1 against Rh2 and 12β-O-Glc-PPD and thus both Rh2 and 12β-O-Glc-PPD could be rapidly transformed to 3β,12β-Di-O-Glc-PPD (Table 3).

A large amount of squalene was accumulated in IN-B, which might be caused by the overexpression of tHMG1 (Table 6). The yield of 3β,12β-Di-O-Glc-PPD is expected to be further improved by overexpression of ERG1 and other downstream enzymes involved in the biosynthetic pathway of 3β,12β-Di-O-Glc-PPD through introduction of increased copies of their genes or use of stronger promoters in the future.

Table 6 Production of 3β,12β-Di-O-Glc-PPD and other precursors in the engineered strains.

Biomass or yield[a]	**IN-A**	**IN-B**	**IN-C**
DCW(g/L)	3.59±0.07	3.15±0.04	2.48±0.11
3β,12β-Di-O-Glc-PPD (mg/L)	6.17±0.35	9.05±0.25	0
DM (mg/L)	1.38±0.14	4.57±0.45	18.25±0.09
PPD (mg/L)	5.04±0.06	11.45±1.56	5.74±0.09

SQ (mg/L)	4.09±0.32	56.93±1.39	6.04±0.12
LST (mg/L)	4.18±0.38	2.98±0.35	14.27±0.28
ER (mg/L)	2.94±0.73	6.16±0.42	2.36±0.20
Total dammarenediol triterpenoids[b] (mg/L)	12.59±0.61	25.08±1.26	23.99±0.36

[a] All given data in this table represents the averages from three independent experiments with corresponding standard deviations. Yield refers to the sum of intracellular and extracellular contents. PPD, protopanaxadiol; DM, dammarenediol-II; SQ, squalene; LST, lanosterol; ER, ergosterol.

[b] Total dammarenediol triterpenoids include DM, PPD and 3β,12β-Di-O-Glc-PPD.

The successful biosynthesis of unnatural ginsenoside 3β,12β-Di-O-Glc-PPD in the metabolically engineered yeasts further confirms that UGT109A1 can catalyze the glycosylation of C3-OH and C12-OH of PPD simultaneously both in vitro and in vivo. The yeast strains engineered in this work can serve as the basis for creating an alternative way for production of 3β,12β-Di-O-Glc-PPD.

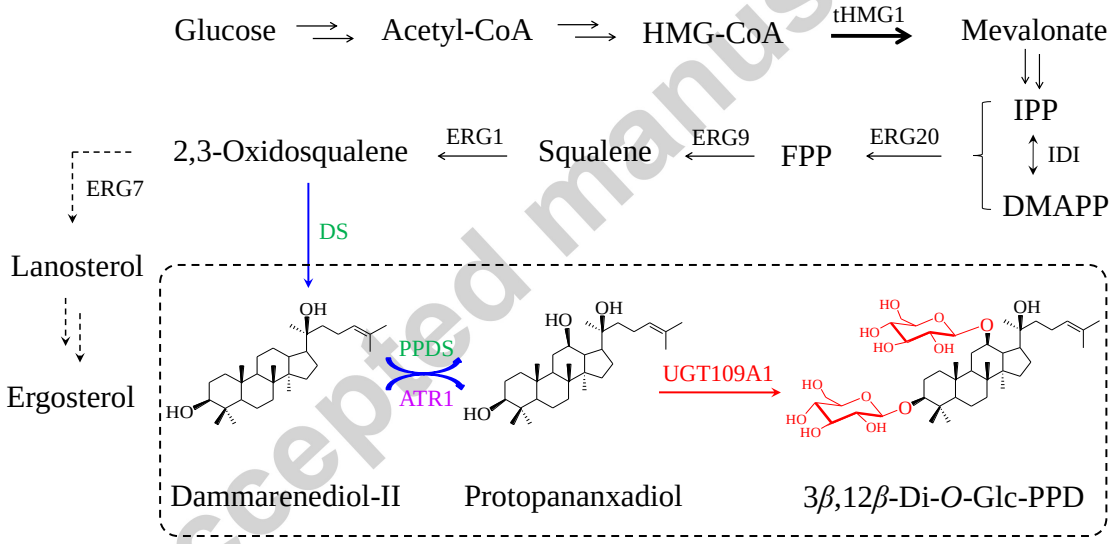


Figure 3. Biosynthetic pathway of 3β,12β-Di-O-Glc-PPD in metabolically engineered yeasts.

Black arrows indicate the intrinsic S. cerevisiae pathway; Multiple-step reactions are shown by multifold arrows; Up-regulated reaction is shown by bold black arrow; Down-regulated reactions are shown by dashed arrows. Blue arrows represent the P. ginseng pathway. Red arrow represents the glycosylation reaction in this study. Black-colored genes are from S. cerevisiae, green-colored genes are from P. ginseng, the purple-colored gene is from A. thaliana, the red-colored gene is the newly

characterized UGT109A1 from B. subtilis.

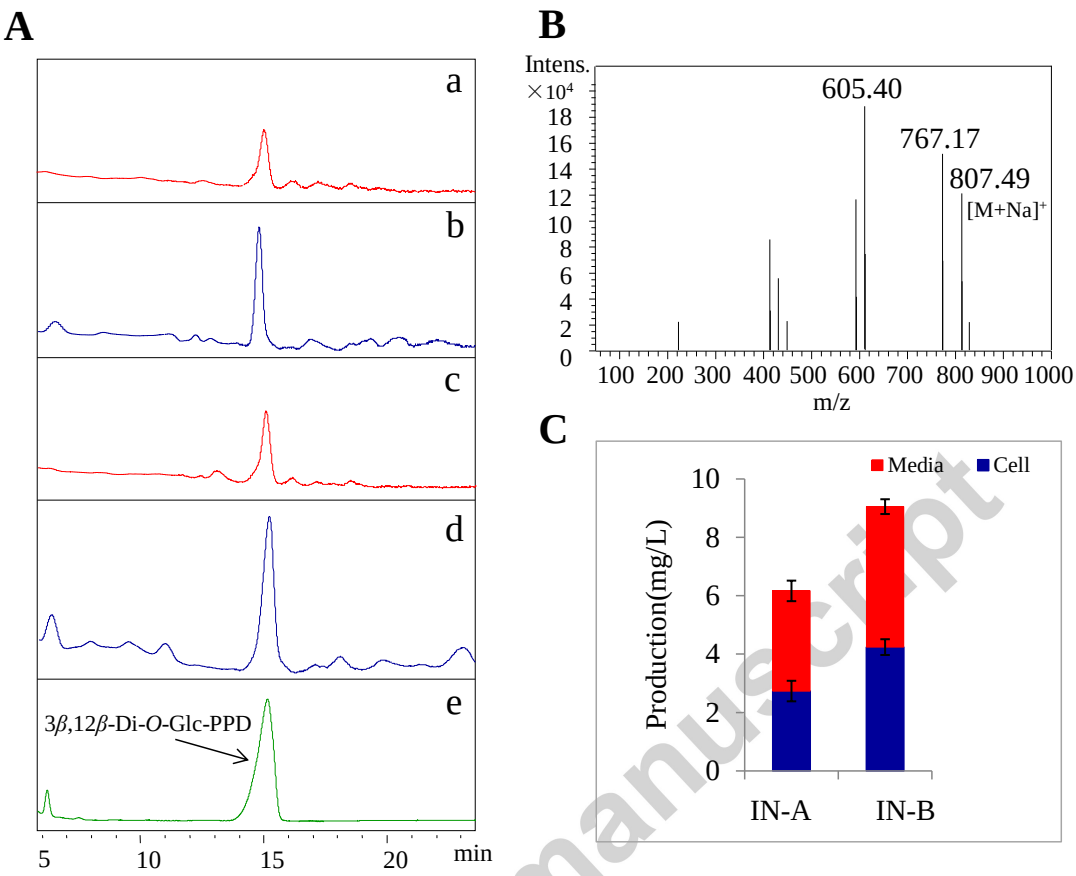


Figure 4. Production of 3β,12β-Di-O-Glc-PPD in metabolically engineered yeasts. (A) HPLC analysis of the products of the recombinant strains extracted by n-butanol. (a) The extracellular production of strain IN-A; (b) The intracellular production of strain IN-A; (c) The extracellular production of strain IN-B; (d) The intracellular production of strain IN-B; (B) The MS spectrum of 3β,12β-Di-O-Glc-PPD produced by the recombinant strains; (C) The yield of 3β,12β-Di-O-Glc-PPD in IN-A and IN-B. Blue and red bars represent intracellular and extracellular production of 3β,12β-Di-O-Glc-PPD respectively.

4. Conclusion

In summary, we identified a new UDP-glycosyltransferase UGT109A1 from B.

subtilis, which transfers a glucose moiety to C3-OH and C20-OH of DM, C3-OH and C12-OH of PPD and PPT, producing 3β-O-Glc-DM, 3β,20S-Di-O-Glc-DM, 3β,12β-Di-O-Glc-PPD and 3β,12β-Di-O-Glc-PPT, respectively. To the best of our knowledge, this is the first report of versatility of UDP-glycosyltransferase to catalyze C3-OH, C12-OH and C20-OH glycosylation of ginsenoside aglycons. Among the unnatural ginsenosides produced by UGT109A1 catalysis, 3β,12β-Di-O-Glc-PPD exhibited higher anti-lung cancer activity than natural ginsenoside Rg3 both in vitro and in vivo. Therefore, metabolically engineered yeasts to produce 3β,12β-Di-O-Glc-PPD were constructed by introducing DS (P. ginseng), PPDS (P. ginseng), ATR1 (A. thaliana) and UGT109A1 (B. subtilis) in S. cerevisiae. The yield of 3β,12β-Di-O-Glc-PPD was increased from 6.17 mg/L to 9.05 mg/L by increasing supplies of 2,3-oxidosqualene through overexpressing tHMG1. It is expected that the yield of 3β,12β-Di-O-Glc-PPD may be significantly increased by further pathway engineering and optimization of fermentation conditions, as proved in the case of taxadiene (Engels et al., 2008), artemisinic acid (Paddon et al., 2013) and PPD (Dai et al., 2013; Zhao et al., 2016). This study may pave the way for the production of 3β,12β-Di-O-Glc-PPD through microbial fermentation in the future, which provides a promising candidate for anti-lung cancer drug discovery.

Acknowledgment

This research was financially supported by the grants of the National Natural Science Foundation of China (81673341), the Beijing Natural Science Foundation (7122115) and CAMS Innovation Fund for Medical Sciences (CIFMS) (2016-I2M-3-012).

References

- Atopkina, L.N., Denisenko, V.A., Uvarova, N.I., Elyakov, G.B., 1988. Semisynthetic analogues of ginsenosides, glycosides from ginseng. Carbohyd. Res. 177, 101-109.
- Atopkina, L.N., Malinovskaya, G.V., Elyakov, G.B., Uvarova, N.I., Woerdenbag, H.J., 1999. Cytotoxicity of natural ginseng glycosides and semisynthetic analogues. Planta Med. 65, 30-34.
- Barbe, V., Cruveiller, S., Kunst, F., Lenoble, P., Meurice, G., Sekowska, A., Vallenet, D., Wang, T., Moszer, I., Medigue, C., Danchin, A., 2009. From a consortium sequence to a unified sequence: the Bacillus subtilis 168 reference genome a decade later. Microbiology 155, 1758-1775.
- Chen, C.F., Chiou, W.F., Zhang J.T., 2008. Comparison of the pharmacological effects of Panax ginseng and Panax quinquefolium. Acta. Pharmacol. Sin. 29, 1103-1108.
- Cheng, H., Zhao, R., Chen, T., Yu, W., Wang, F., Cheng, K., Zhu, P., 2013. Cloning and characterization of the glycoside hydrolases that remove xylosyl groups from 7-β-xylosyl-10-deacetyltaxol and its analogues. Mol. Cell. Proteomics 12, 2236-2248.
- Christensen, L.P., 2009. Ginsenosides chemistry, biosynthesis, analysis, and potential health effects. Adv. Food Nutr. Res. 55, 1-99.
- Dai, Z., Liu, Y., Huang, L., Zhang, X., 2012. Production of miltiradiene by metabolically engineered Saccharomyces cerevisiae. Biotechnol. Bioeng. 109, 2845-2853.
- Dai, Z., Liu, Y., Zhang, X., Shi, M., Wang, B., Wang, D., Huang, L., Zhang, X., 2013. Metabolic engineering of Saccharomyces cerevisiae for production of ginsenosides. Metab. Eng. 20, 146-156.
- Dai, Z., Wang, B., Liu, Y., Shi, M., Wang, D., Zhang, X., Liu, T., Huang, L., Zhang, X., 2014.

Producing aglycons of ginsenosides in bakers' yeast. Sci. Rep. 4, 3698-3698.

- Engels, B., Dahm, P., Jennewein, S., 2008. Metabolic engineering of taxadiene biosynthesis in yeast as a first step towards Taxol (Paclitaxel) production. Metab. Eng. 10, 201-206.
- Gietz, R.D., Schiestl, R.H., Willems, A.R., Willems, A.R., Woods, R.A., 1995. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. Yeast. 11, 355-360.
- Gurung, R.B., Kim, E.H., Oh, T.J., Sohng, J.K., 2013. Enzymatic synthesis of apigenin glucosides by glucosyltransferase (YjiC) from Bacillus licheniformis DSM 13. Mol. Cells 36, 355-361.
- Han, J.Y., Hwang, H.S., Choi, S.W., Kim, H.J., Choi, Y.E., 2012. Cytochrome P450 CYP716A53v2 catalyzes the formation of protopanaxatriol from protopanaxadiol during ginsenoside biosynthesis in Panax ginseng. Plant Cell Physiol. 53, 1535-1545.
- Han, J.Y., Kim, H.J., Kwon, Y.S., Choi, Y.E., 2011. The Cyt P450 Enzyme CYP716A47 catalyzes the formation of protopanaxadiol from dammarenediol-II during ginsenoside biosynthesis in Panax ginseng. Plant cell physiol. 52, 2062-2073.
- Han, J.Y., Kwon, Y.S., Yang, D.C., Jung, Y.R., Choi, Y.E., 2006. Expression and RNA interference-induced silencing of the dammarenediol synthase gene in Panax ginseng. Plant Cell Physiol. 47, 1653-1662.
- Jae, H.K., Bong, G.K., Ahn, J.H., 2006. Glycosylation of flavonoids with a glycosyltransferase from Bacillus cereus. FEMS Microbiol. Lett. 258, 263-268.
- Jung, S.C., Kim, W., Park, S.C., Jeong, J., Park, M.K., Lim, S., Lee, Y., Im, W.T., Lee, J.H., Choi, G., Kim, S.C., 2014. Two ginseng UDP-glycosyltransferases synthesize ginsenoside Rg3 and Rd. Plant Cell Physiol. 55, 2177-2188.

- Kim, Y.J., Zhang, D., Yang, D.C., 2005. Biosynthesis and biotechnological production of ginsenosides. Biotechnol. Adv. 33, 717-735.
- Liang, H., Gao, L., Hu, Z., Gong, T., Chen, J., Yang, J., Zhu, P., 2016. Expression, subcellular localization and characterization of dammarenediol-II synthase of Panax ginseng in Saccharomyces cerevisiae. Acta. Pharm. Sin. 51, 998-1003.
- Luo, S.L., Dang, L.Z., Li, J.F., Zou, C.G., Zhang, K.Q., Li, G.H., 2013. Biotransformation of saponins by endophytes isolated from Panax notoginseng. Chem. Biodivers. 10, 2021-2031.
- Luo, S.L., Dang, L.Z., Zhang, K.Q., Liang, L.M., Li, G.H., 2015. Cloning and heterologous expression of UDP-glycosyltransferase genes from Bacillus subtilis and its application in the glycosylation of ginsenoside Rh1. Lett. Appl. Microbiol. 60, 72-78.
- Paddon, C.J., Westfall, P.J., Pitera, D.J., 2013. High-level semi-synthetic production of the potent antimalarial artemisinin. Nature 496, 528-532.
- Pandey, R.P., Gurung, R.B., Parajuli, P., Koirala, N., Tuoi, L.T., Sohng, J.K., 2014. Assessing acceptor substrate promiscuity of YjiC-mediated glycosylation toward flavonoids. Carbohyd. Res. 393, 26-31.
- Pandey, R.P., Parajuli, P., Shin, J.Y., Lee, J., Lee, S., Hong, Y., Park, Y., Kim, J.S., Sohng, J.K., 2014. Enzymatic biosynthesis of novel resveratrol glucoside and glycoside derivatives. Appl. Environ. Microbiol. 80, 7235-7243.
- Parajuli, P., Pandey, R.P., Koirala, N., Yoon, Y.J., Kim B.G., Sohng J.K., 2014. Enzymatic synthesis of epothilone A glycosides. AMB Express. 4, 31-40.
- Ro, D.K., Paradise, E.M., Ouellet, M., Fisher, K.J., Newman, K.L., Ndungu, J.M., Ho, K.A., Eachus, R.A., Ham, T.S., Kirby, J., Chang, M.C.Y., Withers, S.T., Shiba, Y., Sarpong, R.,

Keasling, J.D., 2006. Production of the antimalarial drug precursor artemisinic acid in engineered yeast. Nature 440, 940-943.

- Shibata, S., 2001. Chemistry and tumor preventing activities of ginseng saponins and some related triterpenoid compounds. J. of Korean Med. Sci. 16, 28-37.
- Shin, B.K., Kwon, S.W., Park, J.H., 2015. Chemical diversity of ginseng saponins from Panax ginseng. J. of Ginseng Res. 16, 287-298.
- Shinkai, K., Akedo, H., Mukai, M., Imamura, F., Isoai, A., Kobayashi, M., Kitagawa, I., 1996. Inhibition of in vitro tumor cell invasion by ginsenoside Rg3. Jpn. J. Cancer Res. 87, 357-362.
- Tansakul, P., Shibuya, M., Kushiro, T., Ebizuka, Y., 2006. Dammarenediol-II synthase, the first dedicated enzyme for ginsenoside biosynthesis, in Panax ginseng. FEBS Lett. 580, 5143-5149.
- Wang, J., Gao, W., Zhang, J., Zuo, B.M., Zhang, L.M., Huang, L.Q., 2012. Advances in study of ginsenoside biosynthesis pathway in Panax ginseng C. A. Meyer. Acta Physiol. Plant. 34, 397-403.
- Wang, P., Wei, Y., Fan, Y., Liu, Q., Wei, W., Yang, C., Zhang, L., Zhao, G., Yue, J., Yan, X., Zhou, Z., 2015. Production of bioactive ginsenosides Rh2 and Rg3 by metabolically engineered yeasts. Metab. Eng. 29, 97-105.
- Wang, Q., Gao, L., Liang, H., Du, G., Gong, T., Yang, J., Zhu, P., 2015. Downregulation of lanosterol synthase gene expression by antisense RNA technology in Saccharomyces cerevisiae. Acta pharm. Sin. 50, 118-122.
- Wei, W., Wang, P., Wei, Y., Liu, Q., Yang, C., Zhao, G., Yue, J., Yan, X., Zhou, Z., 2015. Characterization of Panax ginseng UDP-glycosyltransferases catalyzing protopanaxatriol and biosyntheses of bioactive ginsenosides F1 and Rh1 in metabolically engineered yeasts. Mol.

Plant 8, 1412-1424.

- Xue, J., Liu, Z., Hu, J., Chen, H., Zhang, J., Chen, N., 2006. Ginsenoside Rb1 promotes neurotransmitter release by modulating phosphorylation of synapsins through a cAMP-dependent protein kinase pathway. Brain Res. 1106, 91-98.
- Yang, W., Ying, H., Wu, W., Ye, M., Guo, D., 2014. Saponins in the genus Panax L. (Araliaceae): a systematic review of their chemical diversity. Phytochemistry 106, 7-24.
- Yan, X., Fan, Y., Wei, W., Wang, p., Liu, Q., Wei, Y., Zhang, L., Zhao, G., Yue, J., Zhou, Z., 2014. Production of bioactive ginsenoside compound K in metabolically engineered yeast. Cell Res. 24, 770-773.
- Yu, J., Sun, J., Niu, Y., Li, R., Liao, J., Zhang, F., Yu, B., 2013. Synthetic access toward the diverse ginsenosides. Chem. Sci. 4, 3899-3905.
- Yun, T.K., 2001. Panax ginseng—a non-organ-specific tumor preventive?. Lancet Oncol. 2, 49-55.
- Zhang, Q., Kang, X., Zhao, W., 2006. Antiangiogenic effect of low-dose cyclophosphamide combined with ginsenoside Rg3 on Lewis lung carcinoma. Biochem. Biophs. Res. Commun. 342, 824-828.
- Zhao, F., Bai, P., Liu, T., Li, D., Zhang, X., Lu, W., Yuan, Y., 2016. Optimization of a cytochrome P450 oxidation system for enhancing protopanaxadiol production in Saccharomyces cerevisiae. Biotechnol. Bioeng. 113, 1787-1795.
- Zhuang, Y., Yang, G., Chen, X., Liu, Q., Zhang, X., Deng, Z., Feng, Y., 2017. Biosynthesis of plant-derived ginsenoside Rh2 in yeast via repurposing a key promiscuous microbial enzyme. Metab. Eng. 42, 25-32.

Highlights

- A new UDP-glycosyltransferase BsUGT1 was identified from Bacillus subtilis.Several unnatural ginsenosides were produced by BsUGT1 catalysis.
- 3 β ,12 β -Di-O-Glc-PPD exhibited higher anti-lung cancer activity than Rg3.
- Engineered yeasts were constructed to produce 3 β ,12 β -Di-O-Glc-PPD.